

PATENT ABSTRACTS OF JAPAN

(11)Publication number : 09-044466

(43)Date of publication of application : 14.02.1997

(51)Int.Cl.

G06F 15/18
G06F 17/10

(21)Application number : 07-197461

(71)Applicant : NRI & NCC CO LTD

(22)Date of filing : 02.08.1995

(72)Inventor : SOMUGIYATSUTO
TANKITSUTOWANITSUCHI
TERASAKI TAKESHI
TERAOKA MASAYA

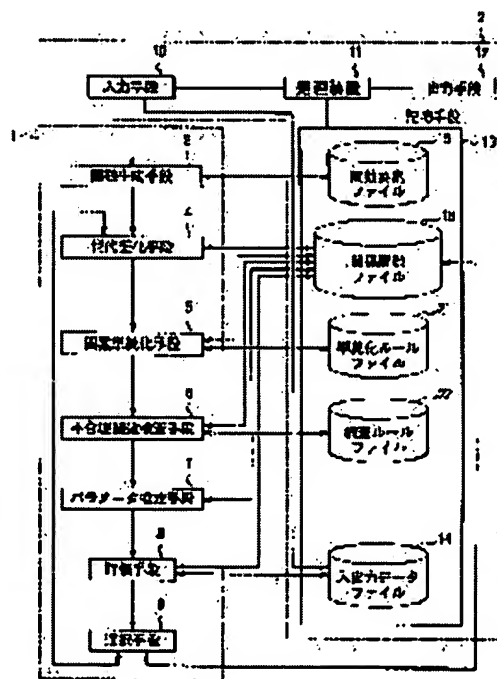
(54) RELATIONAL FUNCTION SEARCHING DEVICE AND METHOD THEREFOR

(57)Abstract:

PROBLEM TO BE SOLVED: To provide a relational function searching device quickly and exactly determining the relational function denoting the relationship between given input data and output data by little computer resources and method therefor.

SOLUTION: This device is provided with a function generation means 3 combining a function element and a parameter and generating a relational function candidate, a generation evolution means 4 evolving a present relational function group to the relational function of the next generation by the evolution technique by a genetic algorithm, a parameter estimation means 7 estimating the optimum value of the parameter of each relational function evolved by the generation evolution

means 4 by a linear regression technique or a non-linear regression technique, an evaluation means 8 evaluating the fitness for given input/output data for the relational function in which the optimum value is estimated by the parameter estimation means 7 and a selection means 9 preferentially selecting the relational function evaluated that the fitness is high by the evaluation means 8.



LEGAL STATUS

[Date of request for examination]	02.08.1995
[Date of sending the examiner's decision of rejection]	06.04.2001
[Kind of final disposal of application other than the examiner's decision of rejection or application converted registration]	
[Date of final disposal for application]	
[Patent number]	3287738
[Date of registration]	15.03.2002
[Number of appeal against examiner's decision of rejection]	2001-07449
[Date of requesting appeal against examiner's decision of rejection]	07.05.2001
[Date of extinction of right]	

Copyright (C); 1998,2003 Japan Patent Office

* NOTICES *

JPO and NCIPI are not responsible for any damages caused by the use of this translation.

- 1.This document has been translated by computer. So the translation may not reflect the original precisely.
- 2.**** shows the word which can not be translated.
- 3.In the drawings, any words are not translated.

CLAIMS

[Claim(s)]

[Claim 1] By the technique of a function generation means to generate a related function candidate combining a functional element and a parameter, and evolution by the genetic algorithm The generation evolution means who evolves the ensemble of a current related function to the ensemble of a next-generation related function, A parameter estimation means to presume the optimum value of the parameter of each relation function which evolved with said generation evolution means by the linear-regression technique or the non-linear regression technique, As opposed to the related function the optimum value was presumed to be by said parameter estimation means Related function retrieval equipment characterized by having an evaluation means to evaluate the fitness to a given I / O data, and a selection means to choose preferentially the related function estimated that fitness is high by said evaluation means.

[Claim 2] Related function retrieval equipment of the publication according to claim 1 characterized by having the function simplification means which rewrites the related function generated by said generation evolution means in a simple form synonymous in formula.

[Claim 3] Related function retrieval equipment according to claim 1 characterized by having a contradiction function inspection means to inspect the form as a formula of a related function, to sort out a related function irrational as a formula, and to except from the object of presumption of a parameter.

[Claim 4] Said generation evolution means is related function retrieval equipment according to claim 1 to 3 characterized by evolving a related function by decussation, mutation, and at least one approach of evolution out of balance.

[Claim 5] said parameter estimation means -- the approach of linear regression, and Guass-Newton -- law and Levenberg-Marquart -- law and Steepest Descent -- law and Quasi-Newton -- law and Simplex Law and Weighting law and Ridge Regression -- law and Full Newton Related function retrieval equipment according to claim 1 to 3 characterized by presuming the optimum value of each relation function candidate's parameter by either of the law.

[Claim 6] Down stream processing which generates the population of a related function candidate's beginning combining a functional element and a parameter, and the related function candidate of said population by the technique of evolution by the genetic algorithm Down stream processing which is evolved and generates a next-generation related function candidate's ensemble, and down stream processing which presumes the optimum value of each of said evolved relation function candidate's parameter by the linear-regression technique or the non-linear regression technique, Down stream processing which evaluates the fitness to a given I / O data to the related function candidate by whom said optimum value was presumed, Down stream processing which chooses the high related function candidate of fitness preferentially among the related function candidates who evaluated said fitness, Said selected, related function candidate's ensemble as a new related function candidate's population Down stream processing which repeats many processings from evolution by said genetic algorithm to selection of the high related function candidate of fitness, The related function retrieval approach characterized by having down stream processing which outputs the related function candidate who reached more than

predetermined fitness, or the related function candidate who has the highest fitness after generation evolution of the count of predetermined as optimal related function.

[Claim 7] The related function retrieval approach according to claim 6 characterized by having down stream processing which rewrites a related function candidate in a simple form synonymous in formula after down stream processing which generates a next-generation related function candidate's ensemble by evolution by said genetic algorithm.

[Claim 8] The related function retrieval approach of claim 6 characterized by having down stream processing which inspects the form as a formula of a related function, sorts out a related function candidate irrational as a formula, and is excepted from the object of presumption of a parameter by evolution by said genetic algorithm after down stream processing which generates a next-generation related function candidate's ensemble.

[Claim 9] Down stream processing of evolution by said genetic algorithm is the related function retrieval approach according to claim 6 to 8 characterized by evolving a related function candidate by decussation, mutation, and at least one approach of an advance out of balance.

[Claim 10] down stream processing of presumption of said parameter -- the approach of linear regression, and Guass-Newton -- law and Levenberg-Marquart -- law and Steepest Descent -- law and Quasi-Newton -- law and Simplex Law and Weighting law and Ridge Regression -- law and Full Newton The related function retrieval approach according to claim 6 to 8 characterized by presuming the optimum value of each relation function candidate's parameter by either of the law.

[Translation done.]

* NOTICES *

JPO and NCIPi are not responsible for any damages caused by the use of this translation.

1. This document has been translated by computer. So the translation may not reflect the original precisely.
2. **** shows the word which can not be translated.
3. In the drawings, any words are not translated.

DETAILED DESCRIPTION

[Detailed Description of the Invention]

[0001]

[Field of the Invention] This invention relates to the related function retrieval equipment which searches for the related function (function showing input/output relation) with which are most satisfied of the relation of those input data groups and output-data groups autonomously, and its approach, when the input data group and the output-data group are given beforehand.

[0002]

[Description of the Prior Art] The relation of the input data and the output data which gave input data and output data (two or more sets of cases are included) as a sample to the computer by development of computer technology in recent years, and were given by computer is learned, and the technique which outputs the output data which balance the relation to new input data based on the learned relation is developed. A neural network is mentioned as such a technique.

[0003] With a neural network, the count unit of a multi-input-output is combined in the shape of a nerve net by the directivity link with weight, and a network is constituted. Each count unit computes one output value to an input value, processing is ****(ed) to other count units through a link with weight for it, and the count unit of the last layer outputs final output data.

[0004] The output value outputted to this neural network's behavior, i.e., a fixed input value, is defined by the gestalt of association of that link, the weight of each link, and the function of a count unit.

[0005] When the sample of input data and output data is given to a neural network as a target of study, a neural network outputs the output data which have processed the input data with the weight of initial value. Next, a neural network compares these output data with the output data of a learning objective, asks for total of that error and the square error of a usual output value and desired value, and he adjusts the weight of each link so that this error may become the smallest. Processing of this study is called "error reverse *****", and can be processed as a problem of the non-linear regression which makes weight of each link a parameter.

[0006] thus, the neural network who learned can output output data also to new input data in the form well approximated to the relation of the I / O data which boiled previously and was given as a learning objective.

[0007] however, although the output data which the above-mentioned neural network approximated the relation of the I / O data given as a learning objective, and balanced it to new input data could be outputted, the learned input/output relation has not been grasped from the outside (input/output relation "is not a positive" about this -- "it being said [-like / positive table] that it is not" again).

[0008] Since input/output relation was not positive table-like, a neural network's approach was not able to evaluate the validity of the input/output relation from a logical viewpoint. That is, in a neural network, the input/output relation has taken what kind of form (form of a function), and what the learned input/output relation can estimate numerically how many relation of the I / O data of a learning objective as was not able to presume whether it would be logically probable.

[0009] For this reason, when approximation and prediction were actually carried out using a neural

network in fields, such as economy, finance, management, marketing, and science, that result approximated or predicted could not presume whether it was the right, but had become the hindrance to utilization.

[0010] On the other hand, the statistical approximation-of-function method for searching for the input/output relation of input data and output data in positive table is learned to the above-mentioned neural network. Generally as this statistical approximation, the linearity multiple regression, the nonlinear multiple regression, etc. are known.

[0011] A linearity multiple regression is the approach of defining the parameters a, b, and c which suit input data and output data most with this linear function as input/output relation supposing a linear function like $Y=aX_1+bX_2+c$.

[0012] On the other hand, a nonlinear multiple regression is the approach of extending the above-mentioned linearity multiple regression further, and defining Parameters a, --, d as a related function supposing a nonlinear function like $Y=aX_1^2+bX_1+cX_2+d$.

[0013] However, since the describing [above] statistical approximation-of-function method had the form with the fixed function which gives input/output relation, it was not able to satisfy suitable input/output relation except the above-mentioned linear function and a nonlinear function.

[0014] The approach of asking for the related function of a free form from the describing [above] statistical approximation-of-function method is John Koza of Stanford University. It is advocated by the professor. It was what generates the related function which suited the input/output relation which predetermined carried out number generation of the related function at random first, and Professor Koza evolved from these related functions repeatedly by the approach of evolution by the genetic algorithm, and was finally given.

[0015] The approach of evolution by the above-mentioned genetic algorithm deals with two or more related functions which are recognizing current existence first as a ream of a series of components like the DNA array of a chromosome, respectively, and makes some elements of those related functions exchange mutually (these are called "decussation"), or throws a new element into some elements of a related function by the fixed probability, and generates the ensemble of new (this is called "mutation") input/output relation. Next, the related function candidate who generated is evaluated, what suited the given input/output relation from a related function candidate's ensemble is chosen, and processing of evolution, evaluation, and selection is further repeated about the selected, related function candidate's ensemble. Thus, what finally satisfies input/output relation is obtained by performing processing of repeat evolution, evaluation, and selection.

[0016] It is called "evolution of time cost" by the approach of evolution by the genetic algorithm to generate the ensemble of other input/output relation from the ensemble of one related function by the above-mentioned evolution, evaluation, and selection. The input/output relation which was made to random by repeating evolution of a generation at the beginning according to the approach of evolution by the genetic algorithm changes the form, while what satisfies the given input/output relation more is chosen, what does not satisfy input/output relation is screened and the related function with which are most satisfied of the input/output relation finally given is generated.

[0017]

[Problem(s) to be Solved by the Invention] However, by the approach of the related function retrieval by the genetic algorithm by above-mentioned Professor Koza, the convergency to the optimal function generation was low, and was not almost able to present a practical application.

[0018] That is, all of the numeric value of parameters, such as a function (these are hereafter called functional element collectively) used as elements, such as four operations, an exponential function, a logarithmic function, and a trigonometric function, a multiplier by which these functional elements are multiplied, and a constant to apply, were chosen as random, and the function consisted of approaches by Professor Koza. For this reason, the countless function existed by difference of a parameter also about the form of one fixed function. The convergency to the form of a function where it was suitable for I/O conditions under the effect of this parameter was very low.

[0019] Moreover, in order to make the function which has the genius which satisfies I/O conditions into

the candidate of the function of each generation who makes it evolve exist, the number of each generation's function had to be made [many]. For this reason, the computer to process had to be what has the very big capacity and throughput.

[0020]

[Means for Solving the Problem] In order to solve the above-mentioned technical problem, the related function retrieval equipment concerning this application claim 1 By the technique of a function generation means to generate a related function candidate combining a functional element and a parameter, and evolution by the genetic algorithm The generation evolution means who evolves the ensemble of a current related function to the ensemble of a next-generation related function, A parameter estimation means to presume the optimum value of the parameter of each relation function which evolved with said generation evolution means by the linear-regression technique or the non-linear regression technique, As opposed to the related function the optimum value was presumed to be by said parameter estimation means It is characterized by having an evaluation means to evaluate the fitness to a given I / O data, and a selection means to choose preferentially the related function estimated that fitness is high by said evaluation means.

[0021] The related function retrieval equipment concerning this application claim 2 is characterized by having the function simplification means which rewrites the related function generated by said generation evolution means in a simple form synonymous in formula in the equipment of claim 1.

[0022] The related function retrieval equipment concerning this application claim 3 is characterized by having a contradiction function inspection means to inspect the form as a formula of a related function, to sort out a related function irrational as a formula, and to except from the object of presumption of a parameter in the equipment of claim 1.

[0023] Said generation evolution means is characterized by the related function retrieval equipment concerning this application claim 4 evolving a related function in claim 1 thru/or one related function retrieval equipment of 3 by decussation, mutation, and at least one approach of evolution out of balance.

[0024] the related function retrieval equipment concerning this application claim 5 -- claim 1 thru/or one related function retrieval equipment of 3 -- setting -- said parameter estimation means -- the approach of linear regression, and Guass-Newton -- law and Levenberg-Marquart -- law and Steepest Descent -- law and Quasi-Newton -- law and Simplex Law and Weighting law and Ridge Regression -- law and Full Newton It is characterized by presuming the optimum value of each relation function candidate's parameter by either of the law.

[0025] The related function retrieval approach concerning this application claim 6 down stream processing which generates the population of a related function candidate's beginning combining a functional element and a parameter, and the related function candidate of said population by the technique of evolution by the genetic algorithm Down stream processing which is evolved and generates a next-generation related function candidate's ensemble, and down stream processing which presumes the optimum value of each of said evolved relation function candidate's parameter by the linear-regression technique or the non-linear regression technique, Down stream processing which evaluates the fitness to a given I / O data to the related function candidate by whom said optimum value was presumed, Down stream processing which chooses the high related function candidate of fitness preferentially among the related function candidates who evaluated said fitness, Said selected, related function candidate's ensemble as a new related function candidate's population Down stream processing which repeats many processings from evolution by said genetic algorithm to selection of the high related function candidate of fitness, It is characterized by having down stream processing which outputs the related function candidate who reached more than predetermined fitness, or the related function candidate who has the highest fitness after generation evolution of the count of predetermined as optimal related function.

[0026] The related function retrieval approach concerning this application claim 7 is characterized by having down stream processing which rewrites a related function candidate in a simple form synonymous in formula after down stream processing which generates a next-generation related function candidate's ensemble by evolution by said genetic algorithm in the approach of claim 6.

[0027] The related function retrieval approach concerning this application claim 8 is characterized by having down stream processing which inspects the form as a formula of a related function, sorts out a related function candidate irrational as a formula, and is excepted from the object of presumption of a parameter by evolution by said genetic algorithm after down stream processing which generates a next-generation related function candidate's ensemble in the approach of claim 6.

[0028] The related function retrieval approach concerning this application claim 9 is characterized by down stream processing of evolution by said genetic algorithm evolving a related function candidate by decussation, mutation, and at least one approach of an advance out of balance in the approach of claim 6 thru/or either of 8.

[0029] the related function retrieval approach concerning this application claim 10 -- the approach of claim 6 thru/or either of 8 -- setting -- down stream processing of presumption of said parameter -- the approach of linear regression, and Guass-Newton -- law and Levenberg-Marquart -- law and Steepest Descent -- law and Quasi-Newton -- law and Simplex Law and Weighting law and Ridge Regression -- law and Full Newton It is characterized by presuming the optimum value of each relation function candidate's parameter by either of the law.

[0030]

[Embodiment of the Invention] One gestalt of operation of this invention is explained using an attached drawing below. Drawing 1 shows the example of 1 configuration and the flow of processing of the related function retrieval equipment by this invention. The related function retrieval equipment 1 of this operation gestalt is used combining it with the external information processor 2. Related function retrieval equipment 1 consists of the function generation means 3, the generation evolution means 4, the function simplification means 5, the contradiction function inspection means 6, the parameter estimation means 7, an evaluation means 8, and a selection means 9. On the other hand, an information processor 2 consists of the input means 10, a processor 11, an output means 12, and a storage means 13. Here, the related function retrieval equipment by this invention is used as the independent equipment combined with an external information processor like the above-mentioned operation gestalt, and also it can also be considered as the configuration which equipped itself with the input means, the output means, the storage means, etc.

[0031] The related function retrieval equipment of this invention is also realizable by starting still more preferably the software which specified actuation of equipment in the information processor of a general configuration of having had the processor, the input means, the storage means, and the output means.

[0032] Next, processing of the related function retrieval equipment of this operation gestalt is explained below. With the related function retrieval equipment 1 of the above-mentioned operation gestalt, input data and output data (this is collectively called I / O data) to ask for the relation between an input and an output first are stored in the I/O data file 14 of the storage means 13 with the input means 10 of an information processor 2. Allocation of the conditions of evolution by the genetic algorithm at the time of asking for a related function (function showing the relation of a I / O data), i.e., the number of the candidates of a related function who perform evolution, (this is called ensemble size), the crossing probability (this is called rate of decussation), a mutation rate, the number (this is called generation number) which repeats evolution, the allowed value of fitness, and the probability at the time of selection etc. inputs into coincidence with an input means 10.

[0033] By the input of the conditions of evolution of the above-mentioned genetic algorithm, the function generation means 3 generates many candidates of a related function based on these conditions. The these-generated related function candidate forms the population of the related function candidate for evolution.

[0034] Here, the approach of generation of the related function which serves as a candidate is explained. The function generation means 3 takes out the functional element of a predetermined individual from the functional-element file 15 of the storage means 13 of an information processor 2, and arranges these functional elements, a variable, and a parameter at random. The above-mentioned functional element also contains operators, such as ** (+) besides functions, such as BEKI ** (^n), sin, cos, log and exp, and --, a decrease (-), ** (*), and ** (/). The above-mentioned variable is a part which substitutes input

data, and only the degree of input data exists. Finally, a multiplier, a constant, etc. substitute a concrete numeric value and the above-mentioned parameter expresses them with notations, such as a, b, c, and --, in the phase of function generation.

[0035] For example, when a functional element, a variable, and a parameter array are set to $((X1) ^)$, (a), (-), (b), (*) (X1), (*) (X2), (+), (c), (*) (X2), (^), (d), (-), and (e), This array means the following relation function candidate $Y=X1^a-bX1 X2+cX2^d-e$. () Inside serves as this related function candidate's component (a function, a variable, and a parameter are included).

[0036] Similarly, when the array of a functional element, a variable, and a parameter becomes $((X1) +)$, (b), and (*) (sin) ((c), (*), and (X2)), this is following related function candidate $Y=X1+bsin (cX2)$. It means.

[0037] Thus, all related function candidates are expressed by the linear array of components, such as a functional element, a variable, and a parameter. Therefore, the function generation means 3 can generate the candidate of a related function by arranging to linearity, combining a functional element, a variable, and a parameter at random.

[0038] By the above-mentioned approach, the function generation means 3 generates a related function candidate until it reaches the ensemble size beforehand specified with the input means 10, and it stores this in the related function file 16 as a related function candidate's population.

[0039] Next, the generation evolution means 4 takes out the population of the candidate of a related function from the related function file 16, and evolves this by the technique of a genetic algorithm. There are decussation, mutation, and many approaches of evolution out of balance in evolution by the technique of this genetic algorithm. the related function retrieval equipment 1 of this invention -- the above -- evolution by the genetic algorithm may be performed by which approach.

[0040] The genetic manipulation which a chromosome carries out pair [every] coupling to the above "decussation" in heredity of a living thing, exchanges a part of gene for it, and obtains the chromosome of a new gene sequence is imitated. That is, deal with a related function candidate as a series of arrays which consist of components, such as a functional element, a variable, and a parameter, make those related function candidates face a pair every, a part of functional element, variable, and parameter are made to exchange, and the candidate of the related function of a new pair is obtained. A mimetic diagram explains this in more detail using the related function candidate who did [above-mentioned] instantiation.

[0041] Drawing 2 shows the situation of decussation of two related function candidates. It is $Y1 =X1^a-bX1 X2+cX2^d-e$ $Y2 =X1+bsin (cX2)$ about the related function candidate before now and decussation. It carries out. These related function candidates can express to a tree structure as shown in drawing 2 . The articulating point (it is called a node) of a tree shows each functional element. On the other hand, the end leaf part (it is called a leaf) of a tree shows the parameter or the variable.

[0042] Decussation is performed by the node of the arbitration of a tree structure, and the node and leaf below that node cross as one lump in this case. For example, it is $Y1$ as shown in drawing 2 . Node 17 and $Y2$ It is $Y1$ when decussation is performed by the node 18. 17 or less-node $cX2^d$ $Y2$ It is mutually exchanged as 1 lump in sin below a node (cX2), respectively. Consequently, $Y1$ $Y2$ It evolves into following $Y1'$ and following $Y2'$, respectively.

$Y1' =X1^a-bX1 X2+sin(cX2)-e$ $Y2' =X1+bcX2^d$, thus every two related function candidates of all of the population are put together, it crosses and a new related function candidate's ensemble is generated by this. The ensemble of this generated related function candidate calls it "a next-generation ensemble" to the population. This next-generation ensemble's related function candidate is stored in the related function file 16.

[0043] To "decussation" exchanging a related function candidate's component mutually, "mutation" changes each related function candidate's component, and generates a new related function candidate.

[0044] That is, the "mutation" in this equipment is changed by the fixed probability to all the related target function candidate's components. Change is random and transposes each functional element, a variable, and a parameter to an unrelated functional element, a variable, and a parameter. The next-generation ensemble who consists of a new related function candidate is generated by this. The next-

generation ensemble of the related function candidate by this mutation is also stored in the related function file 16.

[0045] Next, the approach of "evolution out of balance" is explained. "Evolution out of balance" generates two or more children's (two [for example,]) related function candidate from one related function candidate, and a child [be / it / just like that of parents] and the child from whom some parents' elements changed are made to be intermingled for these children's related function candidate. Thus, by repeating the generation evolution which the child of different nature from parents mixes, the various individuals to an individual which is completely different from an individual [be / it / just like that of the original parents] are generated, and when the fitness in it chooses a high thing, it can converge on the optimal related function.

[0046] Drawing 3 shows notionally the situation of the above-mentioned evolution out of balance. In drawing 3, one parents' related function candidate generates two children's related function candidate. Each parents' related function candidate is leading in case a child is generated. A chain and lagging The template for copying a child called a chain is also as two.

[0047] leading In case a chain 19 (a downward arrow head shows drawing 3) generates a child, it carries out similar succession of parents' characteristic, it makes this a template, and is the other party's lagging. A chain 20 (a upward arrow head shows drawing) is reproduced, and a child is generated. In drawing 3, a broken line shows the chain of the side reproduced. On the other hand, it is lagging. In case a chain 20 generates a child, it makes Mutation e mix in parents' characteristic (here a related function candidate's component) by the fixed probability. The form containing this mutation e is used as a template, and it is the other party's leading. A chain 19 is reproduced and the child from whom parents and a characteristic differed is generated.

[0048] By repeating the above-mentioned evolution out of balance, the various individuals which consist of the individual which carried out similar succession of the first parents' characteristic, an individual which mutation accumulates and is extremely different from the first parents, and an individual of the middle characteristic are generable. Drawing 3 shows the case where three generations evolve, and two mutation e is lagging by one-generation evolution. Three-kind four individuals are generated by the 3rd generation from what does not contain Mutation e at all to whether four mutation e is included and a thing which becomes and is different noting that it mixes in a chain 20.

[0049] The approach of evolution of each genetic algorithm of the above "decussation", "mutation", and "evolution out of balance" may be used independently, and can also be combined and used. For example, after carrying out mutation or carrying out out-of-balance evolution by the probability for it to be fixed after crossing, it can carry out producing mutation in a fixed probability to the whole related function candidate etc.

[0050] Next, it returns to drawing 1 about processing after it of the related function candidate who evolved with the generation evolution means 4, and explains below.

[0051] The function simplification means 5 rewrites these in the function of the simple form of homonymy in formula, taking out the related function candidate of the above-mentioned next generation from the related function file 16, and referring to the simplification Ruhr file 21, after a related function candidate's next-generation ensemble is generated from a related function candidate's population.

[0052] For the generation evolution means 4, since decussation etc. completely makes the component of a related function random, this is $X(a+b+c)$ 1 in a part of related function. $abcX1$ A part [like] may exist (it is here and, for a, b, and c, a parameter and $X1$ are a variable). Since these function parts can be rewritten in a simple form, the function simplification means 5 is $X(a+b+c)$ 1. $dX1$ (new parameter with which d consists of $d=a+b+c$), and $abcX1$ It rewrites to $dX1$ (new parameter with which d consists of $d=abxc$). The parameter which should be dealt with by this can decrease and many processings of presumption of the parameter described henceforth, evaluation, selection, etc. can be performed more easily. Much simplification Ruhr of the form after being simplified with the decision conditions which can simplify a formula is stored in the simplification Ruhr file 21.

[0053] Thus, when it judges and corresponds with reference to the simplification Ruhr of the simplification Ruhr file 21 about each relation function candidate, the function simplification means 5 is

stored in the related function file 16 after it rewrites a related function candidate.

[0054] Next, the contradiction function inspection means 6 sorts out a related function irrational as a formula, taking out the related function candidate rewritten by the simple form from the related function file 16, and referring to the inspection Ruhr file 22 with the function simplification means 5, and presumption of the parameter described later, evaluation, etc. make it the object of processing.

[0055] A thing irrational as a formula is one half of log (negative number) and (negative numbers). ** etc. is needed. The contradiction function inspection means 6 eliminates a related function candidate irrational as a formula, builds the ensemble who consists only of a rational related function candidate, and stores him in the related function file 16.

[0056] Next, the ensemble of the related function candidate by whom the parameter estimation means 7 was sorted out and stored with the above-mentioned contradiction function inspection means 6 is taken out from the related function file 16, and the optimal numeric value of each relation function candidate's parameter is presumed concretely.

[0057] Now and input data x_1, \dots, x_n Output data y_1, \dots, y_n If the related function which shows relation is set to f , it will depend on vector $\theta = [\text{whether the related function } f \text{ approximates input/output relation}]$ ($\theta_1, \dots, \theta_k$) of the combination of the numeric value of a parameter, i.e., a parameter.

[0058] In this case, retrieval of the optimal related function is transposed to the retrieval in parameter space. Retrieval of this optimal parameter is performed so that total of a square error with the related function $f(x_i, \theta)$ of the vector θ of output data y_i ($i = 1, \dots, n$) and a predetermined parameter may be made into min. Here, total of the square error of output data y_i ($i = 1, \dots, n$) and the related function $f(x_i, \theta)$ can be expressed in writing like a bottom type.

$$S^2(\theta) = \sum_{i=1}^n (y_i - f(x_i, \theta))^2$$

Very small modification is added in the direction in which the above $S^2(\theta)$ decreases, and it is made to specifically result in the vector θ of a parameter in it at the optimal related function in the vector of a successive optimal parameter, i.e., the form of the function.

[0059] Gauss-Newton well-known as presumption of the above-mentioned parameter, i.e., the approach of retrieval of the parameter vector in parameter space, -- law -- The Levenberg-Marquart method and Steepest Descent Law, The Quasi-Newton method and Simplex Law and Weighting Law, The Ridge Regression method, Dennis and Gay and Welsch Full Newton boiled and depended Either of the law It can use (refer to Sen Srivastava, Regression Analysis Theory by Springer Verlag work, and Methods and Applications).

[0060] For example, according to the desirable Steepest Descent method, the partial differential of the function S is carried out with each parameter, and a parameter is changed in the gradient direction which is the maximum dive direction. Namely, $\theta(r+1) = \theta(r) - \alpha(r) \Delta \theta(r)$ As $\Delta \theta(r)$ shows, θ is changed serially.

[0061] Here, it is $\theta(r)$. θ and $\Delta \theta(r)$ for which the r -th time was asked $\theta(r)$ The maximum dive direction which can be set, θ_1 [i.e.,] of $S^2(\theta)$, --, and θ_k The maximum of the partial differential which can be set, and $\alpha(r)$ $\theta(r)$ The width of face of change in the maximum dive direction is shown.

[0062] With the Steepest Descent method, it is $\alpha(r)$. By considering as a suitable value, by repeating change of θ , $S^2(\theta)$ comes to take the minimum value, presumes the parameter vector θ at this time ($= \theta_1, \dots, \theta_k$) to be an optimal parameter, and determines the concrete form of a related function.

[0063] By the above-mentioned approach, the parameter estimation means 7 presumes the optimal parameter of the candidate of each relation function, and stores it in the related function file 16.

[0064] Next, the evaluation means 8 computes each relation function candidate's fitness with reference to the I/O data file 14 about the candidate of each relation function concretely specified by the parameter

estimation means 7. In this case, fitness can also be considered as the total $S^2(\theta)$ of the square error computed previously, and can give a gravity to the group of a I / O data, and can also give high fitness by the related function candidate who adapts himself to the group of an important I / O data. The related function candidate who was able to give fitness with the evaluation means 8 is stored in the related function file 16.

[0065] Next, the selection means 9 chooses a related function candidate from current related function candidates focusing on what has high fitness.

[0066] There are some which combined the roulette method, elite Conservation Act, and these as the selection approach by the above-mentioned selection means 9.

[0067] The roulette method assigns each candidate the probability proportional to fitness, and chooses the high candidate of fitness by the high probability. According to this approach, possibility that what has the low present fitness may be chosen and that selected thing will evolve into the optimal thing by evolution in the future can be collateralized.

[0068] On the other hand, elite Conservation Act chooses a fixed number of candidates sequentially from the high candidate of fitness. Although its current fitness is low while what has current fitness high [this approach] has the advantage surely chosen, what has good genius also has the fault screened.

[0069] The selection by the selection means 9 of this operation gestalt combines the advantage of the describing [above] roulette method and elite Conservation Act. That is, a fixed number of candidates are chosen as the high order of fitness by elite Conservation Act, and it chooses by the roulette method about what remains. A candidate with high fitness is surely chosen by this, and even if the present fitness is low, it can leave possibility that the candidate from whom fitness will become high will also be chosen by the fixed probability, by evolution.

[0070] The candidate of the related function chosen by the selection means 9 is stored in the related function file 16. The related function candidate with very high fitness which fitness became less than the predetermined reference value at this time is outputted as an adaptation solution in the form where the concrete numeric value was substituted for the parameter. When such an adaptation solution is not found, the selected, related function candidate is stored in the related function file 16 in the form equipped with the parameter with a notation. In this case, in the next generation evolution, the generation evolution by the genetic algorithm is presented with the ensemble of the related function candidate chosen in the former generation as the new population.

[0071] A series of processings of evolution by the above-mentioned genetic algorithm, inspection of the simplification contradiction function of a function, presumption of a parameter, evaluation, and selection are called evolution of the time cost by the genetic algorithm. By repeating the generation evolution by this genetic algorithm, related function retrieval equipment 1 can change a related function candidate in the direction with which are satisfied of input/output relation gradually, and, finally can obtain the related function of the optimal form.

[0072] In addition, each of above-mentioned function simplification and contradiction function inspection is for making presumption of a parameter easy, and can be omitted if needed.

[0073] Below, the effectiveness of the related function retrieval equipment of this invention in the comparison with the conventional related function retrieval equipment by Professor Koza is explained.

[0074] Drawing 4 compares and shows progress of the retrieval to the optimum solution of the related function retrieval equipment by this invention, and the related function retrieval equipment by Professor Koza.

[0075] Although the fitness of the related function candidate by Professor Koza's related function retrieval equipment is in the gradient which increases as a whole as shown in drawing 4, the fitness after each ***** cost is fluctuated violently. On the other hand, it can be increased for every ***** cost by the fitness of the related function candidate of the related function retrieval equipment by this invention, and it can reach the optimum solution of a target slack relation function almost linearly.

[0076] Therefore, also in the count of a repeat of generation evolution, the related function retrieval equipment by this invention can reach the same target by the sharply small count compared with the equipment by Professor Koza.

[0077] That such a phenomenon arises has a cause in generating in the form where the figure concrete into a multiplier or a constant part was substituted for generating a related function candidate with Professor Koza's equipment. Thus, the convergency which decision of a thing which the cause of the lowness of fitness depends on numeric values, such as what is depended on the form of a function, and a multiplier, with the related function of the form where the concrete figure was substituted for the multiplier or the constant cannot be performed, but tends toward a target was bad, and retrieval of a related function was almost impossible in the throughput of the usual computer.

[0078] On the other hand, with the related function retrieval equipment by this invention, the form and parameter of a function are separated and the numeric value of the optimal parameter is calculated in the form of each relation function focusing on evolution of the form of a related function. For this reason, the candidate of a related function can approach the form of the quickly optimal related function by evolution, and can determine a related function in the form where the concrete numeric value was finally substituted for the parameter. The computer which thereby actually has a usual capacity and a usual throughput can also search for a related function in the rational range.

[0079]

[Effect of the Invention] The related function retrieval equipment by this invention constitutes the candidate showing the input/output relation of the given I / O data of a related function with a functional element, a variable, and a parameter notation, evolves the candidate of these related functions by decussation, mutation, evolution out of balance, etc., and presumes a parameter by technique, such as non-linear regression, after an appropriate time so that clearly from the above explanation. Evaluation of a related function candidate is performed to the related function candidate who presumed the parameter, and fitness leaves a next-generation candidate the candidate of a high related function as a function (what expressed the parameter with the notation) of the form. The related function retrieval equipment which searches for quickly the input/output relation of the given I / O data in the form of a positive table-function with the usual throughput and the computer of capacity by this, and its approach are realizable.

[Translation done.]

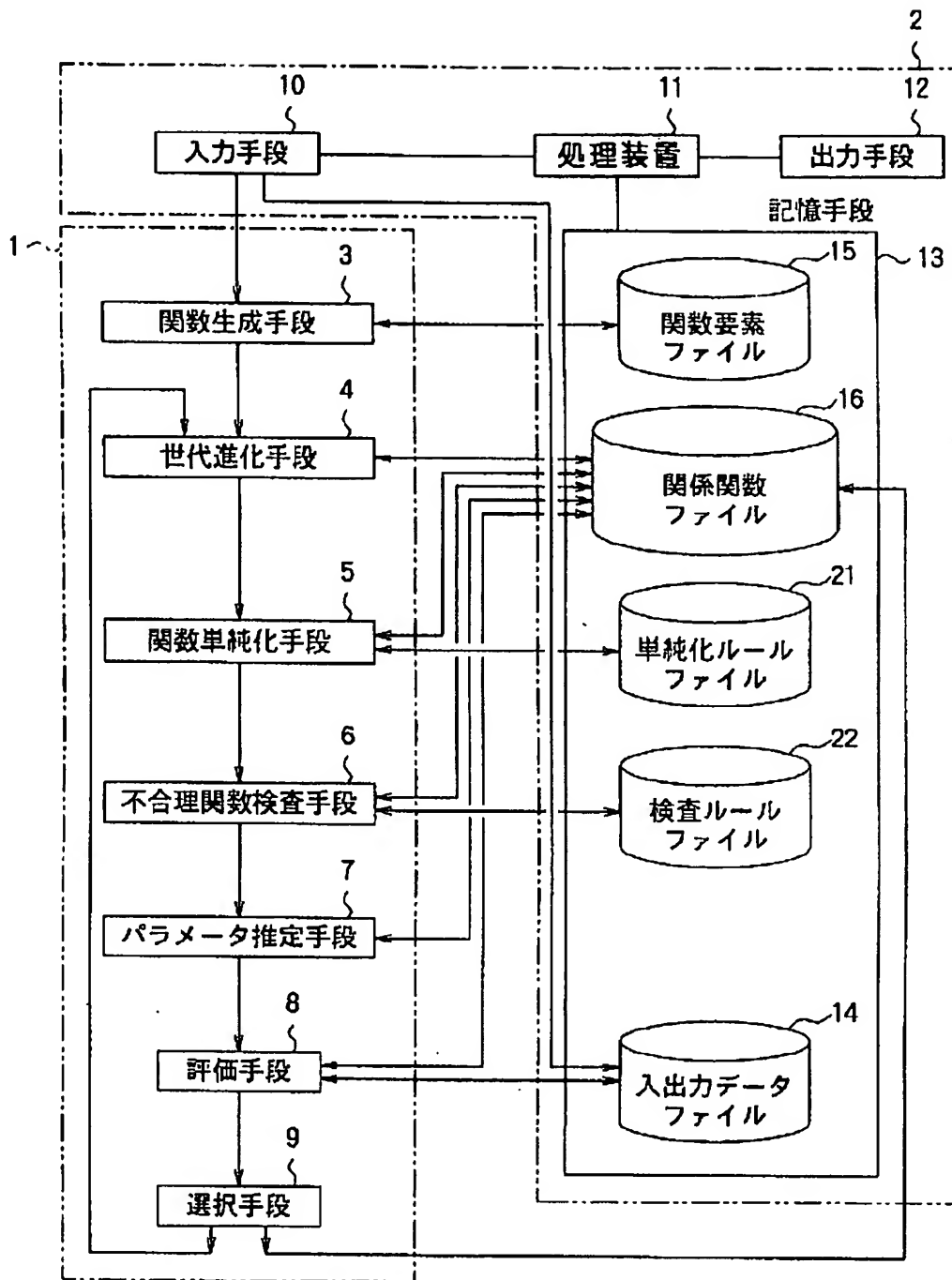
* NOTICES *

JPO and NCIPi are not responsible for any damages caused by the use of this translation.

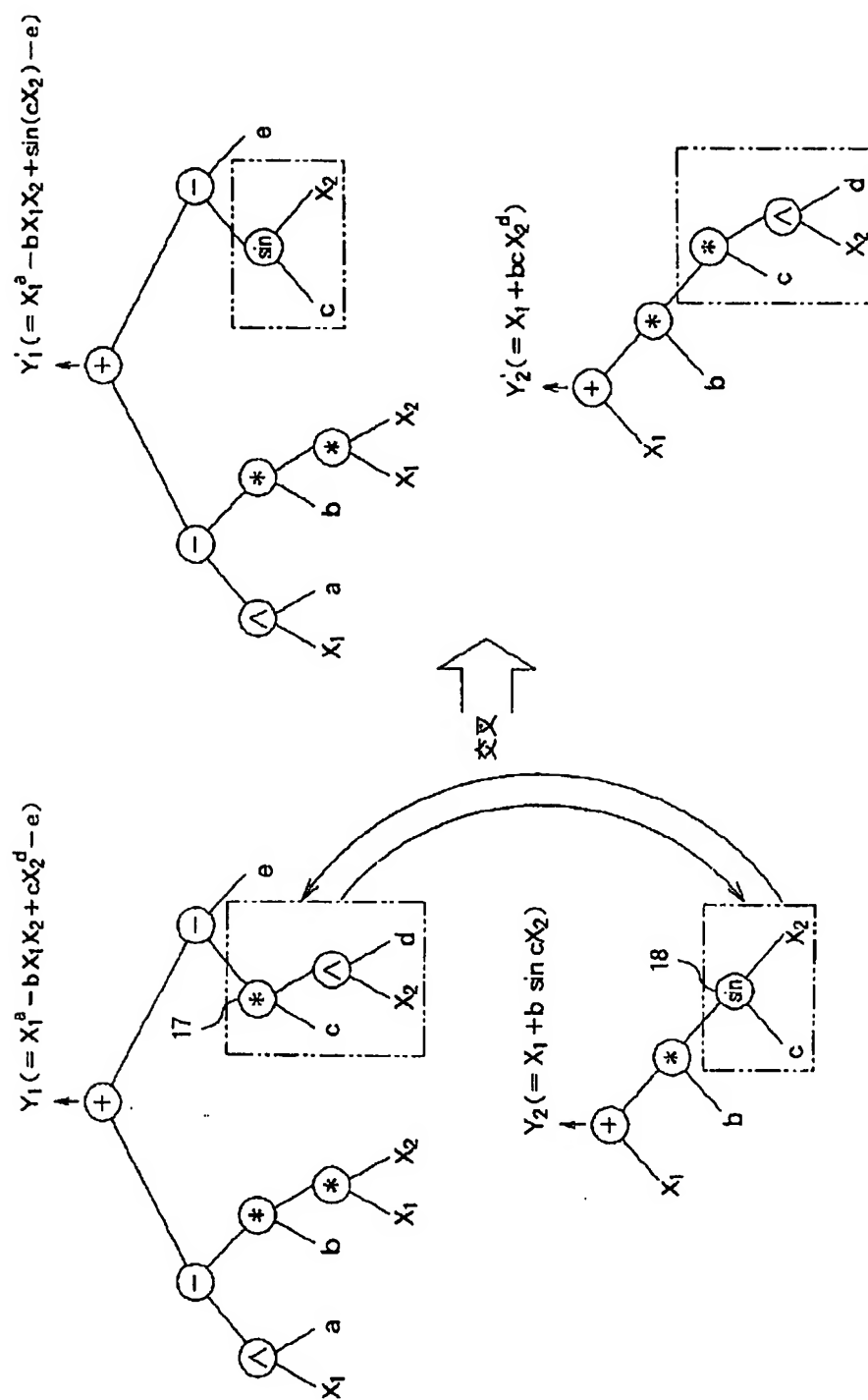
- 1.This document has been translated by computer. So the translation may not reflect the original precisely.
- 2.**** shows the word which can not be translated.
- 3.In the drawings, any words are not translated.

DRAWINGS

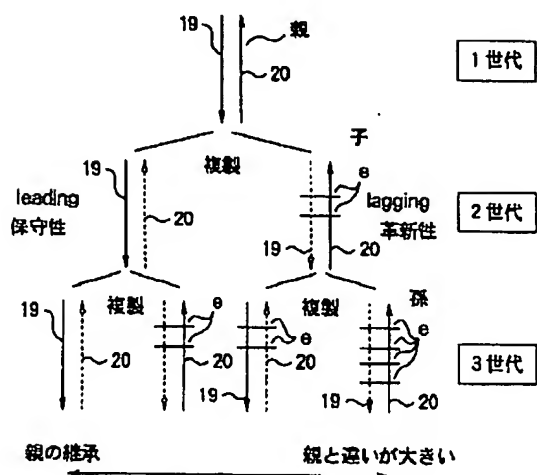
[Drawing 1]



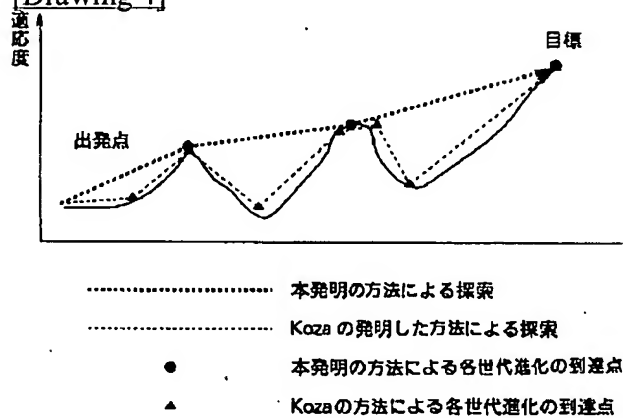
[Drawing 2]



[Drawing 3]



[Drawing 4]



[Translation done.]

(19) 日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11) 特許出願公開番号

特開平9-44466

(43) 公開日 平成9年(1997)2月14日

(51) Int.Cl. ⁹	識別記号	庁内整理番号	F I	技術表示箇所
G 0 6 F 15/18 17/10	5 5 0		G 0 6 F 15/18 15/31	5 5 0 C Z

審査請求 有 請求項の数10 O L (全 11 頁)

(21) 出願番号 特願平7-197461

(22) 出願日 平成7年(1995)8月2日

(71) 出願人 000155469

株式会社野村総合研究所

東京都中央区日本橋1丁目10番1号

(72) 発明者 ソムギヤット、タンキットワニッチ

神奈川県横浜市保土ヶ谷区神戸町134番地

株式会社野村総合研究所内

(72) 発明者 寺 崎 健

神奈川県横浜市保土ヶ谷区神戸町134番地

株式会社野村総合研究所内

(72) 発明者 寺 岡 正 也

神奈川県横浜市保土ヶ谷区神戸町134番地

株式会社野村総合研究所内

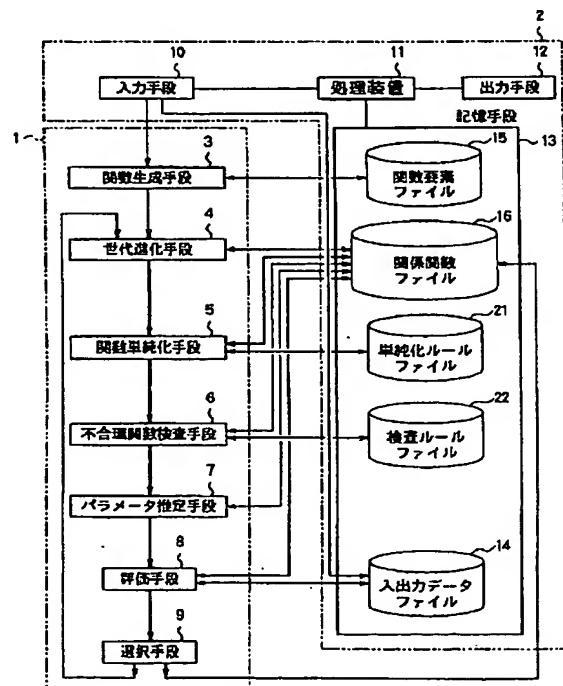
(74) 代理人 弁理士 佐藤 一雄 (外3名)

(54) 【発明の名称】 関係関数探索装置及びその方法

(57) 【要約】

【課題】 所与の入力データと出力データとの関係を示す関係関数を、少ない計算機資源によって迅速かつ的確に求める関係関数探索装置およびその方法を提供する。

【解決手段】 関数要素とパラメータとを組み合わせる関係関数候補を生成する関数生成手段3と、遺伝的アルゴリズムによる進化の手法によって、現在の関係関数の集団を次世代の関係関数の集団に進化させる世代進化手段4と、前記世代進化手段によって進化した各関係関数のパラメータの最適値を、線形回帰手法あるいは非線形回帰手法によって推定するパラメータ推定手段7と、前記パラメータ推定手段によって最適値が推定された関係関数に対して、所与の入出力データに対する適応度を評価する評価手段8と、前記評価手段によって適応度が高いと評価された関係関数を優先的に選択する選択手段9と、を備えた。



【特許請求の範囲】

【請求項 1】関数要素とパラメータとを組み合わせる関係関数候補を生成する関数生成手段と、

遺伝的アルゴリズムによる進化の手法によって、現在の関係関数の集団を次世代の関係関数の集団に進化させる世代進化手段と、

前記世代進化手段によって進化した各関係関数のパラメータの最適値を、線形回帰手法あるいは非線形回帰手法によって推定するパラメータ推定手段と、

前記パラメータ推定手段によって最適値が推定された関係関数に対して、所与の入出力データに対する適応度を評価する評価手段と、

前記評価手段によって適応度が高いと評価された関係関数を優先的に選択する選択手段と、を備えていることを特徴とする関係関数探索装置。

【請求項 2】前記世代進化手段によって生成された関係関数を、数式的に同義な単純な形に書き換える関数単純化手段を備えていることを特徴とする請求項 1 に記載の記載の関係関数探索装置。

【請求項 3】関係関数の数式としての形を検査し、数式として不合理な関係関数を選別してパラメータの推定の対象から除外する不合理関数検査手段を備えていることを特徴とする請求項 1 に記載の関係関数探索装置。

【請求項 4】前記世代進化手段は、交叉、突然変異、不均衡進化の少なくとも一つの方法によって、関係関数を進化させることを特徴とする請求項 1 ないし請求項 3 のいずれかに記載の関係関数探索装置。

【請求項 5】前記パラメータ推定手段は、線形回帰の方法、Guass-Newton法、Levenberg-Marquart法、Steepest Descent法、Quasi-Newton法、Simplex 法、Weighting 法、Ridge Regression法、Full Newton 法のいずれかによって各関係関数候補のパラメータの最適値を推定することを特徴とする請求項 1 ないし請求項 3 のいずれかに記載の関係関数探索装置。

【請求項 6】関数要素とパラメータとを組み合わせる関係関数候補の最初の母集団を生成する処理工程と、前記母集団の関係関数候補を、遺伝的アルゴリズムによる進化の手法により、進化させて次世代の関係関数候補の集団を生成する処理工程と、

前記進化した各関係関数候補のパラメータの最適値を、線形回帰手法あるいは非線形回帰手法によって推定する処理工程と、

前記最適値が推定された関係関数候補に対して所与の入出力データに対する適応度を評価する処理工程と、

前記適応度を評価した関係関数候補のうち適応度の高い関係関数候補を優先的に選択する処理工程と、

前記選択された関係関数候補の集団を新たな関係関数候補の母集団として、前記遺伝的アルゴリズムによる進化から適応度の高い関係関数候補の選択までの諸処理を繰り返す処理工程と、

所定の適応度以上に達した関係関数候補、あるいは所定回数の世代進化後の最も高い適応度を有する関係関数候補を、最適な関係関数として出力する処理工程と、を有していることを特徴とする関係関数探索方法。

【請求項 7】前記遺伝的アルゴリズムによる進化によって次世代の関係関数候補の集団を生成する処理工程の後に、関係関数候補を数式的に同義な単純な形に書き換える処理工程を有していることを特徴とする請求項 6 に記載の関係関数探索方法。

【請求項 8】前記遺伝的アルゴリズムによる進化によって次世代の関係関数候補の集団を生成する処理工程の後に、関係関数の数式としての形を検査し、数式として不合理な関係関数候補を選別してパラメータの推定の対象から除外する処理工程を有していることを特徴とする請求項 6 の関係関数探索方法。

【請求項 9】前記遺伝的アルゴリズムによる進化の処理工程は、交叉、突然変異、不均衡進行の少なくとも一つの方法によって関係関数候補を進化させることを特徴とする請求項 6 ないし請求項 8 のいずれかに記載の関係関数探索方法。

【請求項 10】前記パラメータの推定の処理工程は、線形回帰の方法、Guass-Newton法、Levenberg-Marquart 法、Steepest Descent法、Quasi-Newton法、Simplex 法、Weighting 法、Ridge Regression法、Full Newton 法のいずれかによって各関係関数候補のパラメータの最適値を推定することを特徴とする請求項 6 ないし請求項 8 のいずれかに記載の関係関数探索方法。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】本発明は、予め入力データ群と出力データ群が与えられている場合に、それらの入力データ群と出力データ群の関係を最も満足する関係関数（入出力関係を表す関数）を自律的に探索する関係関数探索装置及びその方法に関する。

【0002】

【従来の技術】近年のコンピュータ技術の発達により、サンプルとして入力データと出力データ（複数組の場合を含む）をコンピュータに与え、コンピュータによって与えられた入力データと出力データの関係を学習し、その学習した関係に基づいて新たな入力データに対してその関係に見合う出力データを出力する技術が開発されている。このような技術としてたとえばニューラルネットワークが挙げられる。

【0003】ニューラルネットワークとは、多入力出力の計算ユニットを、重み付きの方向性リンクによって神経網状に結合してネットワークを構成したものである。各々の計算ユニットは入力値に対して一つの出力値を算出し、それを重み付きのリンクを介して他の計算ユニットに処理を伝播し、最終層の計算ユニットは、最終的な出力データを出力する。

【0004】このニューラルネットワークの挙動、すなわち一定の入力値に対して出力する出力値は、そのリンクの結合の形態、各リンクの重み、計算ユニットの関数によって定められる。

【0005】ニューラルネットワークに学習の目標として入力データと出力データのサンプルを与えた場合、ニューラルネットワークは初期値の重みによって入力データを処理してある出力データを出力する。次にニューラルネットワークは、この出力データと学習目標の出力データとを比較し、その誤差、通常出力値と目標値の二乗誤差の総和を求め、この誤差が最も小さくなるように各リンクの重みを調整する。この学習の処理は、「誤差逆伝播」と呼ばれており、各リンクの重みをパラメータとする非線形回帰の問題として処理することができる。

【0006】このように学習したニューラルネットワークは、新たな入力データに対しても、先に学習目標として与えた入出力データの関係によく近似した形で出力データを出力することができる。

【0007】しかし、上記ニューラルネットワークは、学習目標として与えられた入出力データの関係を近似して新たな入力データに対してそれに見合った出力データを出力することができるが、その学習した入出力関係は外部からは把握することができなかった（このことを、入出力関係が「陽でない」また「陽表的でない」という）。

【0008】ニューラルネットワークの方法は、入出力関係が陽表的でないために、その入出力関係の妥当性を論理的な観点から評価することができなかった。つまり、ニューラルネットワークでは、学習した入出力関係がどの程度学習目標の入出力データの関係を数値的に評価することができるものも、その入出力関係がどのような形（関数の形）を採っていて、それが論理的に確からしいか否かを推定することができなかった。

【0009】このため、経済、金融、経営、マーケティング、自然科学等の分野で実際にニューラルネットワークを用いて近似や予測をする場合、その近似または予測した結果が正しいか否かを推定することができず、実用化への妨げとなっていた。

【0010】一方、上記ニューラルネットワークに対して、入力データと出力データの入出力関係を陽表的に求める統計的な関数近似法が知られている。この統計的近似法としては線形重回帰や非線形重回帰などが一般に知られている。

【0011】線形重回帰は

$$Y = aX_1 + bX_2 + c$$

のような線形関数を入出力関係として想定し、この線形関数で最も入力データと出力データに適合するパラメータ a 、 b 、 c を定める方法である。

【0012】一方、非線形重回帰は、上記線形重回帰をさらに拡張し、

$$Y = aX_1^2 + bX_1 + cX_2 + d$$

のような非線形関数を関係関数として想定してパラメータ a 、 \dots 、 d を定める方法である。

【0013】しかし、上記統計的関数近似法は、入出力関係を与える関数が固定的な形を有しているので、上記線形関数および非線形関数以外の適当な入出力関係を満足することはできなかった。

【0014】上記統計的関数近似法に対して自由な形の関係関数を求める方法がスタンフォード大学の John Koza 教授によって提唱されている。Koza 教授は、最初にランダムに関係関数を所定の個数生成し、これらの関係関数から遺伝的アルゴリズムによる進化の方法によって繰り返し進化させ、最終的に与えられた入出力関係に適合した関係関数を生成するものであった。

【0015】上記遺伝的アルゴリズムによる進化の方法は、最初に現在存在している複数の関係関数をそれぞれ染色体の DNA 配列のように一連の構成要素の連なりとして取扱い、それらの関係関数の要素の一部をたがいに交換させ（これらを“交叉”という）、あるいは一定の確率で関係関数の要素の一部に新たな要素を投入し（これを“突然変異”という）新たな入出力関係の集団を生成する。次に、その生成した関係関数候補を評価し、関係関数候補の集団から、与えられた入出力関係に適合したものを選択し、選択した関係関数候補の集団についてさらに、進化、評価、選択の処理を繰り返す。このように繰り返し進化、評価、選択の処理を行うことによって、最終的に入出力関係を満足するものを得る。

【0016】上記進化、評価、選択によって一つの関係関数の集団から他の入出力関係の集団を生成することを、遺伝的アルゴリズムによる進化の方法では“一世代の進化”という。遺伝的アルゴリズムによる進化の方法によれば、世代の進化を繰り返すことによって、当初ランダムに作った入出力関係がその形を変え、与えられた入出力関係をより満足するものが選択される一方、入出力関係を満足しないものは淘汰され、最終的に与えられた入出力関係をもっとも満足する関係関数が生成される。

【0017】

【発明が解決しようとする課題】しかしながら、上記 Koza 教授による遺伝的アルゴリズムによる関係関数探索の方法では、最適な関数生成への収束性が低く、殆ど実用的な用途に供することができなかった。

【0018】すなわち、Koza 教授による方法では、加減乗除、指数関数、対数関数、三角関数等の要素となる関数（以下、これらをまとめて関数要素という）と、これら関数要素に乗じる係数や加える定数等のパラメータの数値をすべてランダムに選択して関数を構成していた。このため、一つの定まった関数の形についても、パラメータの相違によって無数の関数が存在した。このパラメータの影響によって入出力条件に適した関数の形への収

束性がきわめて低かった。

【0019】また、進化させる各世代の関数の候補の中に入出力条件を満足する素質を有する関数を存在させるために、各世代の関数の個数を多くしなければならなかった。このため、処理するコンピュータはきわめて大きな容量と処理能力を有しているものでなければならなかった。

【0020】

【課題を解決するための手段】上記課題を解決するために、本願請求項1に係る関係関数探索装置は、関数要素とパラメータとを組み合わせる関係関数候補を生成する関数生成手段と、遺伝的アルゴリズムによる進化の手法によって、現在の関係関数の集団を次世代の関係関数の集団に進化させる世代進化手段と、前記世代進化手段によって進化した各関係関数のパラメータの最適値を、線形回帰手法あるいは非線形回帰手法によって推定するパラメータ推定手段と、前記パラメータ推定手段によって最適値が推定された関係関数に対して、所与の入出力データに対する適応度を評価する評価手段と、前記評価手段によって適応度が高いと評価された関係関数を優先的に選択する選択手段と、を備えていることを特徴とするものである。

【0021】本願請求項2に係る関係関数探索装置は、請求項1の装置において、前記世代進化手段によって生成された関係関数を、数式的に同義な単純な形に書き換える関数単純化手段を備えていることを特徴とするものである。

【0022】本願請求項3に係る関係関数探索装置は、請求項1の装置において、関係関数の数式としての形を検査し、数式として不合理な関係関数を選別してパラメータの推定の対象から除外する不合理関数検査手段を備えていることを特徴とするものである。

【0023】本願請求項4に係る関係関数探索装置は、請求項1ないし3のいずれかの関係関数探索装置において、前記世代進化手段は、交叉、突然変異、不均衡進化の少なくとも一つの方法によって、関係関数を進化させることを特徴とするものである。

【0024】本願請求項5に係る関係関数探索装置は、請求項1ないし3のいずれかの関係関数探索装置において、前記パラメータ推定手段は、線形回帰の方法、Guass-Newton法、Levenberg-Marquart法、Steepest Descent法、Quasi-Newton法、Simplex法、Weighting法、Ridge Regression法、Full Newton法のいずれかによって各関係関数候補のパラメータの最適値を推定することを特徴とするものである。

【0025】本願請求項6に係る関係関数探索方法は、関数要素とパラメータとを組み合わせる関係関数候補の最初の母集団を生成する処理工程と、前記母集団の関係関数候補を、遺伝的アルゴリズムによる進化の手法により、進化させて次世代の関係関数候補の集団を生成する

処理工程と、前記進化した各関係関数候補のパラメータの最適値を、線形回帰手法あるいは非線形回帰手法によって推定する処理工程と、前記最適値が推定された関係関数候補に対して所与の入出力データに対する適応度を評価する処理工程と、前記適応度を評価した関係関数候補のうち適応度の高い関係関数候補を優先的に選択する処理工程と、前記選択された関係関数候補の集団を新たな関係関数候補の母集団として、前記遺伝的アルゴリズムによる進化から適応度の高い関係関数候補の選択までの諸処理を繰り返す処理工程と、所定の適応度以上に達した関係関数候補、あるいは所定回数の世代進化後の最も高い適応度を有する関係関数候補を、最適な関係関数として出力する処理工程と、を有していることを特徴とするものである。

【0026】本願請求項7に係る関係関数探索方法は、請求項6の方法において、前記遺伝的アルゴリズムによる進化によって次世代の関係関数候補の集団を生成する処理工程の後に、関係関数候補を数式的に同義な単純な形に書き換える処理工程を有していることを特徴とするものである。

【0027】本願請求項8に係る関係関数探索方法は、請求項6の方法において、前記遺伝的アルゴリズムによる進化によって次世代の関係関数候補の集団を生成する処理工程の後に、関係関数の数式としての形を検査し、数式として不合理な関係関数候補を選別してパラメータの推定の対象から除外する処理工程を有していることを特徴とするものである。

【0028】本願請求項9に係る関係関数探索方法は、請求項6ないし8のいずれかの方法において、前記遺伝的アルゴリズムによる進化の処理工程は、交叉、突然変異、不均衡進行の少なくとも一つの方法によって関係関数候補を進化させることを特徴とするものである。

【0029】本願請求項10に係る関係関数探索方法は、請求項6ないし8のいずれかの方法において、前記パラメータの推定の処理工程は、線形回帰の方法、Guass-Newton法、Levenberg-Marquart法、Steepest Descent法、Quasi-Newton法、Simplex法、Weighting法、Ridge Regression法、Full Newton法のいずれかによって各関係関数候補のパラメータの最適値を推定することを特徴とするものである。

【0030】

【発明の実施の形態】以下に本発明の実施の一形態について添付の図面を用いて説明する。図1は、本発明による関係関数探索装置の一構成例とその処理の流れを示している。本実施形態の関係関数探索装置1は、外部の情報処理装置2と組み合わされて使用される。関係関数探索装置1は、関数生成手段3と、世代進化手段4と、関数単純化手段5と、不合理関数検査手段6と、パラメータ推定手段7と、評価手段8と、選択手段9とからなる。一方、情報処理装置2は、入力手段10と、処理装

置 1 1 と、出力手段 1 2 と、記憶手段 1 3 とからなる。ここで、本発明による関係関数探索装置は、上記実施形態のように外部の情報処理装置と組み合わせられる独立の装置とする他、それ自体に手段、出力手段、記憶手段等を備えた構成とすることもできる。

【0031】さらに好ましくは、処理装置と手段と記憶手段と出力手段とを備えた一般的な構成の情報処理装置において、装置の動作を規定したソフトウェアを起動することにより、本発明の関係関数探索装置を実現することもできる。

【0032】次に本実施形態の関係関数探索装置の処理について以下に説明する。上記実施形態の関係関数探索装置 1 では、最初に入力と出力の関係を求めたいデータと出力データ（これをまとめて入出力データという）を、情報処理装置 2 の手段 1 0 によって記憶手段 1 3 の入出力データファイル 1 4 に格納する。同時に、手段 1 0 によって、関係関数（入出力データの関係を表す関数）を求める際の遺伝的アルゴリズムによる進化の条件、すなわち進化を行う関係関数の候補の数（これを集団サイズという）、交叉する確率（これを交叉率という）、突然変異率、進化を繰り返す数（これを世代数という）、適応度の許容値、選択時の確率の配分等を入力する。

【0033】上記遺伝的アルゴリズムの進化の条件のにより、関数生成手段 3 は、それら条件に基づいて関係関数の候補を多数生成する。これら生成された関係関数候補は、進化のための関係関数候補の母集団を形成する。

【0034】ここで、候補となる関係関数の生成の方法について説明する。関数生成手段 3 は、情報処理装置 2 の記憶手段 1 3 の関数要素ファイル 1 5 から所定個の関数要素を取り出し、これら関数要素と変数とパラメータとをランダムに配列する。上記関数要素は、べき乗（ \wedge ）、 \sin 、 \cos 、 \log 、 \exp 、…等の関数の他、加（+）、減（-）、乗（*）、除（/）等の演算子も含む。上記変数は、データを代入する部分であり、データの次数だけ存在する。上記パラメータは、係数や定数等、最終的には具体的な数値を代入するものであり、関数生成の段階では a、b、c、…等の記号で表わしたものである。

【0035】たとえば、関数要素と変数とパラメータ配列が (X1) (\wedge) (a) (-) (b) (*) (X1) (*) (X2) (+) (c) (*) (X2) (\wedge) (d) (-) (e) となった場合、この配列は下記の関係関数候補

$$Y = X1^a - b X1 X2 + c X2^d - e$$

を意味している。() 内はこの関係関数候補の構成要素（関数と変数とパラメータとを含む）となる。

【0036】同様にたとえば、関数要素と変数とパラメータの配列が、(X1) (+) (b) (*) (sin) 50

((c) (*) (X2)) となった場合、これは下記の関係関数候補

$$Y = X1 + b \sin(c X2)$$

を意味している。

【0037】このようにすべての関係関数候補は、関数要素と変数とパラメータ等の構成要素の線形配列によって表わされる。したがって、関数生成手段 3 は、関数要素と変数とパラメータとをランダムに組み合わせて線形に配列することによって関係関数の候補を生成することができる。

【0038】上述方法により、関数生成手段 3 は、予め手段 1 0 によって指定した集団サイズに達するまで関係関数候補を生成し、これを関係関数候補の母集団として関係関数ファイル 1 6 に格納する。

【0039】次に、世代進化手段 4 が、関係関数ファイル 1 6 から関係関数の候補の母集団を取り出し、これを遺伝的アルゴリズムの手法によって進化させる。この遺伝的アルゴリズムの手法による進化には、交叉、突然変異、不均衡進化の諸方法がある。本発明の関係関数探索装置 1 では、上記いずれの方法によって遺伝的アルゴリズムによる進化を行ってもよい。

【0040】上記「交叉」とは、生物の遺伝において染色体が一對ずつカップリングし、遺伝子の一部を交換して新たな遺伝子配列の染色体を得る遺伝子操作を真似たものである。すなわち、関係関数候補を関数要素、変数、パラメータ等の構成要素からなる一連の配列として取扱い、それらの関係関数候補を一對ずつ相対させ、関数要素、変数、パラメータの一部を交換させて、新たな一對の関係関数の候補を得るのである。このことを上記例示した関係関数候補を用いて模式図によってさらに詳しく説明する。

【0041】図 2 は、二つの関係関数候補の交叉の様子を示している。今、交叉前の関係関数候補を

$$Y1 = X1^a - b X1 X2 + c X2^d - e$$

$$Y2 = X1 + b \sin(c X2)$$

とする。これらの関係関数候補は、図 2 に示すようなツリー構造に表わすことができる。ツリーの分節点（ノードという）は、各関数要素を示している。一方、ツリーの末葉部分（リーフという）は、パラメータあるいは変数を示している。

【0042】交叉は、ツリー構造の任意のノードで行なわれ、この場合そのノード以下のノードとリーフは一つのかたまりとして交叉される。たとえば、図 2 に示すように Y1 のノード 1 7 と Y2 のノード 1 8 で交叉が行われると、Y1 のノード 1 7 以下の $c X2^d$ と Y2 のノード以下の $\sin(c X2)$ がそれぞれ一かたまりとして互いに交換される。この結果、Y1 と Y2 はそれぞれ、下記の Y1' と Y2' とに進化する。

$$Y1' = X1^a - b X1 X2 + \sin(c X2) - e$$

$$Y2' = X1 + b c X2^d$$

このようにして、母集団の関係関数候補のすべては、二つずつ組み合わせられ、交叉され、これによって新しい関係関数候補の集団が生成される。この生成された関係関数候補の集団は、母集団に対して「次世代集団」という。この次世代集団の関係関数候補は、関係関数ファイル16に格納される。

【0043】「交叉」が関係関数候補の構成要素を互いに交換するのに対して、「突然変異」は個々の関係関数候補の構成要素を変化させて新しい関係関数候補を生成する。

【0044】すなわち、本装置における「突然変異」は、対象となる関係関数候補のすべての構成要素に対して、一定の確率で変化させる。変化はランダムであり、それぞれの関数要素、変数、パラメータを無関係な関数要素、変数、パラメータに置き換える。これによって、新しい関係関数候補からなる次世代集団が生成される。この突然変異による関係関数候補の次世代集団も、関係関数ファイル16に格納される。

【0045】次に「不均衡進化」の方法について説明する。「不均衡進化」は、一つの関係関数候補から複数の（たとえば二つの）子関係関数候補を生成し、これらの子関係関数候補には、親とそっくりの子と、親の要素の一部が変化した子とが混在するようにする。このように、親と異なる資質の子が混入する世代進化を繰り返すことによって、元の親とそっくりの個体から全く異なる個体までの多様な個体が生成され、その中の適応度が高いものを選択することによって最適な関係関数に収束することができる。

【0046】図3は上記不均衡進化の様子を概念的に示したものである。図3において、一つの親の関係関数候補は、二つの子の関係関数候補を生成する。それぞれの親の関係関数候補は、子生成の際にleading 鎖とlagging 鎖という子をコピーするためのひな型を二つもっている。

【0047】leading 鎖19（図3において下向きの矢印によって示す）は、子生成の際に親の形質をそっくり継承し、これをひな型として相手方のlagging 鎖20（図において上向きの矢印によって示す）を複製し、子生成する。複製される側の鎖は、図3において破線で示す。一方、lagging 鎖20は、子生成の際に親の形質（ここでは関係関数候補の構成要素）に一定の確率で突然変異eを混入させる。この突然変異eを含んだ形をひな型にして相手方のleading 鎖19を複製し、親と形質の異なった子を生成する。

【0048】上記不均衡進化を繰り返すことによって、最初の親の形質をそっくり継承した個体と、突然変異が累積して最初の親と極めて異なる個体と、その中間の形質の個体とからなる多様な個体を生成することができる。図3は、3世代進化した場合を示しており、1世代進化で2個の突然変異eがlagging 鎖20に混入すると

して、3世代目には突然変異eを全く含まないものから、4個の突然変異eを含むかなり異なるものまで、3種類4個の個体が生成される。

【0049】上記「交叉」と「突然変異」と「不均衡進化」の各遺伝的アルゴリズムの進化の方法は単独で用いてもよく、また、組み合わせて使用することもできる。たとえば、交叉を行った後に一定の確率で突然変異させ、あるいは不均衡進化させた後に全体の関係関数候補に対して一定の確率で突然変異を生じさせる等することができる。

【0050】次に世代進化手段4によって進化した関係関数候補のそれ以降の処理について図1に戻って以下に説明する。

【0051】関数単純化手段5は、関係関数候補の母集団から関係関数候補の次世代集団が生成された後に、上記次世代の関係関数候補を関係関数ファイル16から取り出し、単純化ルールファイル21を参照しながら、これらを数式的に同義の単純な形の関数に書き換える。

【0052】これは、世代進化手段4は、関係関数の構成要素を全くランダムに交叉等するので、関係関数の一部にたとえば $(a+b+c)X1$ や $abcX1$ のような部分が存在することがある（ここでa, b, cはパラメータ、X1は変数）。これらの関数部分は単純な形に書き換えることができるので、関数単純化手段5は、 $(a+b+c)X1$ を $dX1$ （ d は $d=a+b+c$ からなるあらたなパラメータ）、 $abcX1$ を $dX1$ （ d は $d=a \times b \times c$ からなる新たなパラメータ）に書き換える。これによって取り扱うべきパラメータ等が減少し、以降に述べるパラメータの推定、評価、選択等の諸処理をより簡単に行うことができる。単純化ルールファイル21には、数式を単純化できる判断条件と単純化した後の形の単純化ルールが多数格納されている。

【0053】このようにして関数単純化手段5は、各関係関数候補について単純化ルールファイル21の単純化ルールを参照して判断し、該当する場合には、関係関数候補を書き換えた後に関係関数ファイル16に格納する。

【0054】次に、不合理関数検査手段6は、関数単純化手段5によって単純な形に書き換えられた関係関数候補を関係関数ファイル16から取り出し、検査ルールファイル22を参照しながら数式として不合理な関係関数を選別し、後に述べるパラメータの推定や評価等の処理の対象とする。

【0055】数式として不合理なものとは、たとえばlog（負の数）、（負の数）の1/2乗、等のようなものである。不合理関数検査手段6は、数式として不合理な関係関数候補を排除して、合理的な関係関数候補のみからなる集団をつくって関係関数ファイル16に格納する。

【0056】次にパラメータ推定手段7が上記不合理関数検査手段6によって選別されて格納された関係関数候

補の集団を関係関数ファイル16から取り出し、各関係関数候補のパラメータの最適な数値を具体的に推定する。

【0057】今、入力データ x_1, \dots, x_n と出力データ y_1, \dots, y_n の関係を示す関係関数を f とすると、関係関数 f が入出力関係を近似しているか否かはパラメータの数値の組合せ、すなわちパラメータのベクトル $\theta = (\theta_1, \dots, \theta_k)$ による。

$$S^2(\theta) = \sum_{i=1}^n (y_i - f(x_i, \theta))^2$$

具体的には、上記 $S^2(\theta)$ が減少する方向にパラメータのベクトル θ に微小な変更を加えて、逐次的最適パラメータのベクトルに、すなわちその関数の形における最適な関係関数に至るようにする。

【0059】上記パラメータの推定、すなわちパラメータ空間におけるパラメータベクトルの探索の方法として公知の Gauss-Newton 法、Levenberg-Marquart 法、Steepest Descent 法、Quasi-Newton 法、Simplex 法、Weighting 法、Ridge Regression 法、Dennis, Gay and Welsch による Full Newton 法のいずれかを用いることができる (Sen, Srivastava, Springer Verlag 著による Regression Analysis Theory, Methods and Applications 参照)。

【0060】たとえば、好ましい Steepest Descent 法によれば、関数 S を各パラメータで偏微分して、その最急降下方向である gradient 方向にパラメータを変更する。すなわち、

$$\theta(r+1) = \theta(r) - \alpha(r) \Delta(r)$$

で示すように θ を逐次変化させる。

【0061】ここで、 $\theta(r)$ は r 回目に求めた θ 、 $\Delta(r)$ は $\theta(r)$ における最急降下方向、すなわち、 $S^2(\theta)$ の $\theta_1, \dots, \theta_k$ における偏微分の最大値、 $\alpha(r)$ は $\theta(r)$ の最急降下方向への変化の幅を示している。

【0062】Steepest Descent 法では、 $\alpha(r)$ を適当な値とすることにより、 θ の変化を繰り返すことにより $S^2(\theta)$ は最小値をとるようになり、このときのパラメータベクトル $\theta (= \theta_1, \dots, \theta_k)$ を最適パラメータと推定して関係関数の具体的な形を決定する。

【0063】パラメータ推定手段7は、上記方法によって各関係関数の候補の最適パラメータを推定し、関係関数ファイル16に格納する。

【0064】次に評価手段8は、パラメータ推定手段7によって具体的に特定された各関係関数の候補について、入出力データファイル14を参照し、各関係関数候補の適応度を算出する。この場合、適応度は先に算出した二乗誤差の総和 $S^2(\theta)$ とすることもできるし、また、入出力データの組に軽重を与え、重要な入出力データの組に適応する関係関数候補により高い適応度を与えることもできる。評価手段8によって適応度を与えら

*【0058】この場合、最適な関係関数の探索は、パラメータ空間内の探索に置き換えられる。この最適なパラメータの探索は、出力データ y_i ($i=1, \dots, n$) と所定のパラメータのベクトル θ の関係関数 $f(x_i, \theta)$ との二乗誤差の総和を最小に行う。ここで、出力データ y_i ($i=1, \dots, n$) と関係関数 $f(x_i, \theta)$ の二乗誤差の総和は、下式のように書き表すことができる。

れた関係関数候補は関係関数ファイル16に格納される。

【0065】次に選択手段9は、現在の関係関数候補の中から適応度が高いものを中心に関係関数候補を選択する。

【0066】上記選択手段9による選択方法としてはルーレット法とエリート保存法とこれらを組み合わせたものがある。

【0067】ルーレット法は、適応度に比例した確率を各候補に割り当て、適応度の高い候補を高い確率で選択するようにしたものである。この方法によれば、現在適応度が低いものも選択されることがあり、その選択されたものが将来進化によって最適なものに進化する可能性を担保することができる。

【0068】これに対してエリート保存法は、適応度の高い候補から順に一定の数の候補を選択するものである。この方法は、現在適応度が高いものは必ず選択される利点を有する反面、現在適応度が低いが高素質が良いものは淘汰される欠点も有している。

【0069】本実施形態の選択手段9による選択は、上記ルーレット法とエリート保存法の長所を組み合わせたものである。すなわち、適応度の高い順に候補をエリート保存法によって一定数選択し、残るものについてはルーレット法によって選択するものである。これにより、適応度が高い候補が必ず選択され、また、現在適応度が低くても進化によって適応度が高いものになる候補も一定の確率で選択される可能性を残すことができる。

【0070】選択手段9によって選択された関係関数の候補は、関係関数ファイル16に格納される。このとき適応度が所定の基準値以内となった極めて適応度が高い関係関数候補は、パラメータに具体的な数値を代入した形で適合解として出力される。このような適合解が見つからない場合は、選択された関係関数候補は記号によるパラメータを備えた形で関係関数ファイル16に格納される。この場合、次の世代進化では、前世代で選択された関係関数候補の集団が新たな母集団として遺伝的アルゴリズムによる世代進化に供される。

【0071】上記遺伝的アルゴリズムによる進化、関数の単純化不合理関数の検査、パラメータの推定、評価、

選択の一連の処理を、遺伝的アルゴリズムによる一世代の進化という。関係関数探索装置1は、この遺伝的アルゴリズムによる世代進化を繰り返すことによって、関係関数候補を入出力関係を満足する方向に徐々に変化させ、最終的に最適な形の関係関数を得ることができる。

【0072】なお、上記関数単純化および不合理関数検査は、いずれもパラメータの推定を容易にするためのものであり、必要に応じて省略することができる。

【0073】以下に、Koza教授による従来の関係関数探索装置との比較における本発明の関係関数探索装置の効果について説明する。

【0074】図4は、本発明による関係関数探索装置とKoza教授による関係関数探索装置の最適解への探索の経過を比較して示している。

【0075】図4に示すように、Koza教授の関係関数探索装置による関係関数候補の適応度は全体として増加する傾度にあるものの、各世代進代後の適応度は激しく増減する。これに対して本発明による関係関数探索装置の関係関数候補の適応度は、世代進代ごとに増加し目標たる関係関数の最適解にほぼ直線的に到達することができる。

【0076】したがって世代進化の繰り返し回数においても、本発明による関係関数探索装置は、Koza教授による装置に比べて大幅に少ない回数によって同一の目標に到達することができる。

【0077】このような現象が生じるのは、Koza教授の装置では関係関数候補を生成するのに係数や定数部分に具体的な数字を代入した形で生成していることに原因がある。このように係数や定数に具体的な数字を代入した形の関係関数では、適応度の低さの原因が、関数の形によるものか、係数等の数値によるものかの判断ができず、目標に向っての収束性が悪く、通常の計算機の処理能力では関係関数の探索はほとんど不可能であった。

【0078】これに対して、本発明による関係関数探索装置では、関数の形とパラメータを切り離し、関係関数の形の進化を中心とし、各関係関数の形において最適なパラメータの数値を求めている。このため、関係関数の候補は進化によって急速に最適な関係関数の形に近づき、最終的にはパラメータに具体的な数値を代入した形で関係関数を決定することができる。これにより、実際に通常の容量と処理能力を有する計算機によっても合理的な範囲で関係関数の探索を行うことができる。

【0079】

【発明の効果】以上の説明から明らかなように、本発明*

*による関係関数探索装置は、与えられた入出力データの入出力関係を表わす関係関数の候補を関数要素と変数とパラメータ記号とによって構成し、これらの関係関数の候補を交叉、突然変異、不均衡進化等によって進化させ、しかる後にパラメータを非線形回帰等の手法によって推定する。関係関数候補の評価は、パラメータを推定した関係関数候補に対して行い、適応度が高い関係関数の候補をその形の関数（パラメータを記号によって表わしたもの）として次世代の候補に残してゆく。これにより、通常の処理能力と容量の電子計算機によって、与えられた入出力データの入出力関係を陽表的な関数の形で素早く求める関係関数探索装置及びその方法を実現することができる。

【図面の簡単な説明】

【図1】本発明による関係関数探索装置の一構成例とその処理の流れを示したブロック図。

【図2】「交叉」の様子を概念的に示した図。

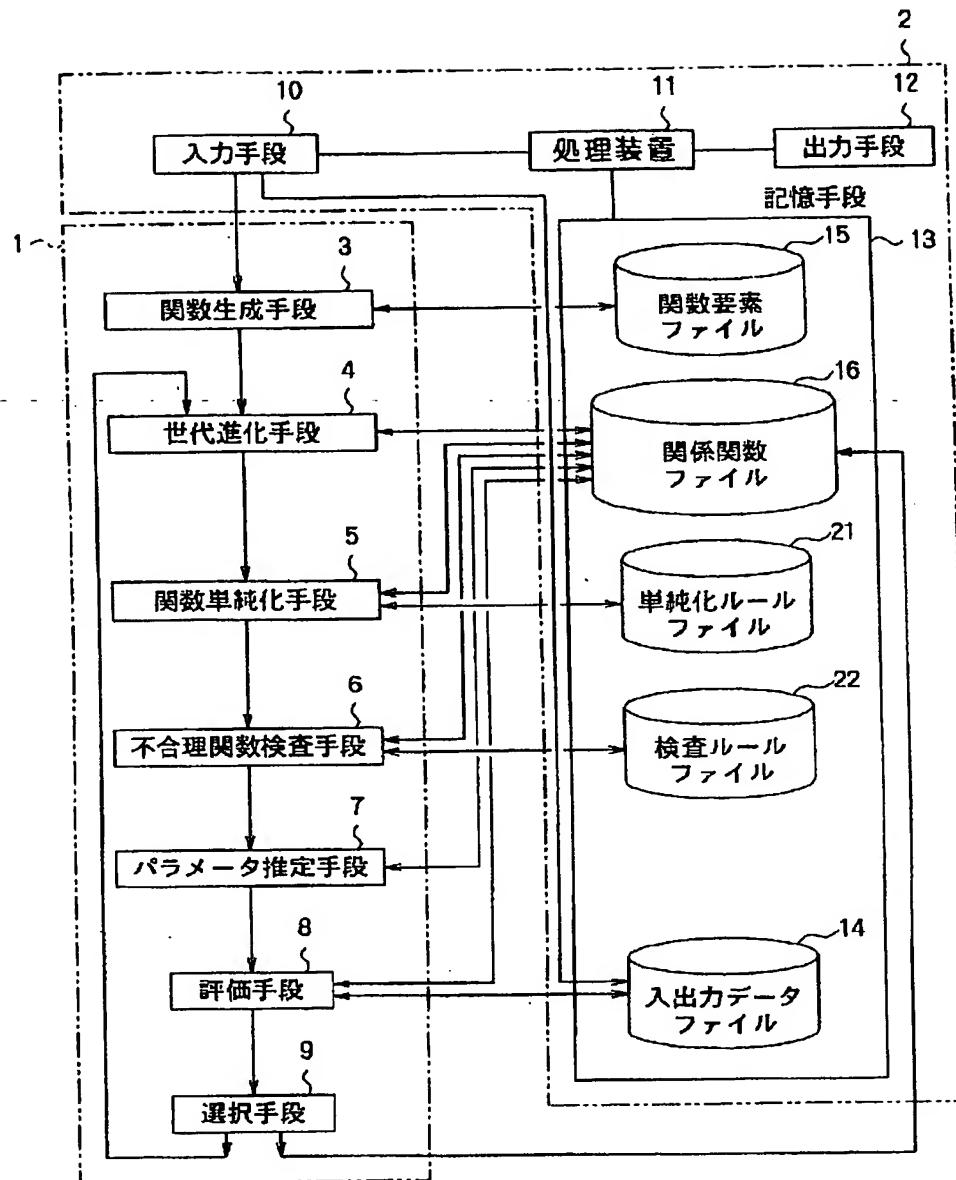
【図3】「不均衡進化」の様子を概念的に示した図。

【図4】本発明による関係関数探索装置とKoza教授による関係関数探索装置の、最適解探索の経過を比較して示したグラフ。

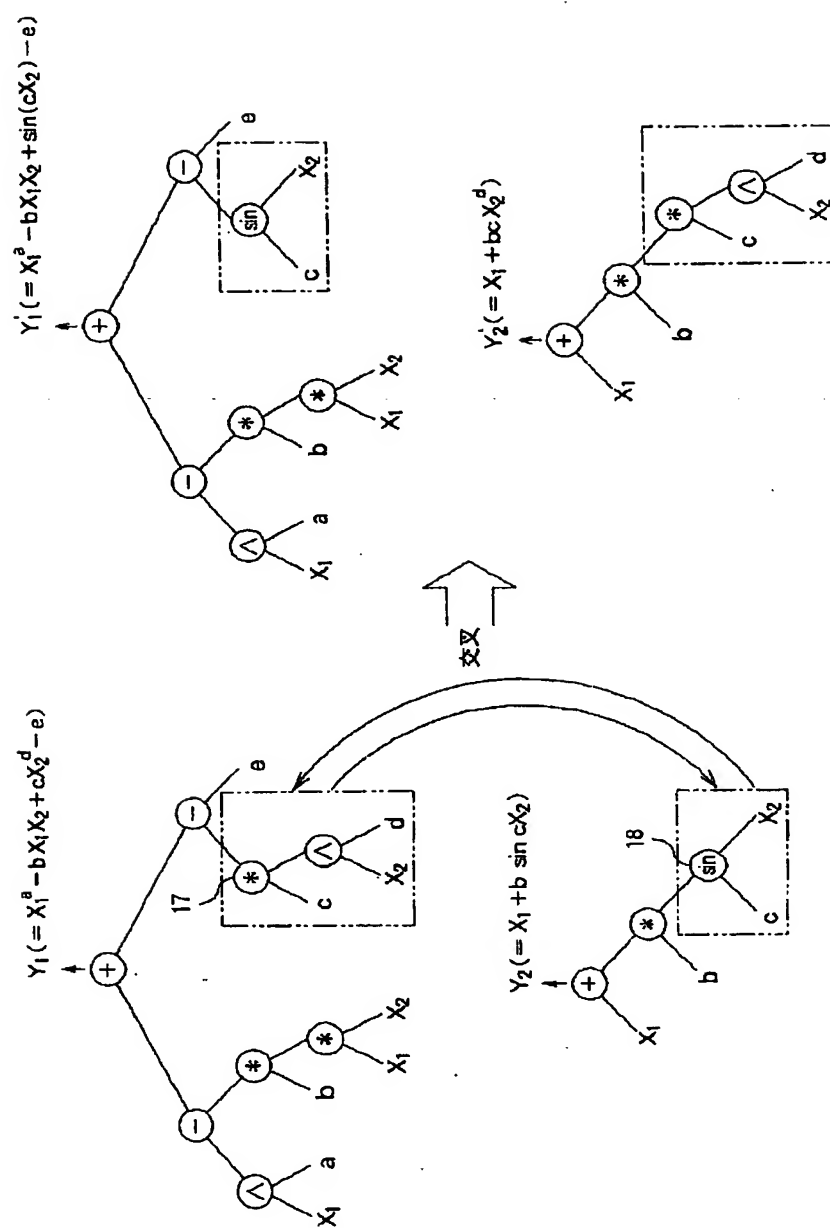
【符号の説明】

- 1 関係関数探索装置
- 2 情報処理装置
- 3 関数生成手段
- 4 世代進化手段
- 5 関数単純化手段
- 6 不合理関数検査手段
- 7 パラメータ推定手段
- 8 評価手段
- 9 選択手段
- 10 入力手段
- 11 処理装置
- 12 出力手段
- 13 記憶手段
- 14 入出力データファイル
- 15 関数要素ファイル
- 16 関係関数ファイル
- 17 ノード
- 18 ノード
- 19 leading鎖
- 20 lagging鎖
- 21 単純化ルールファイル
- 22 検査ルールファイル

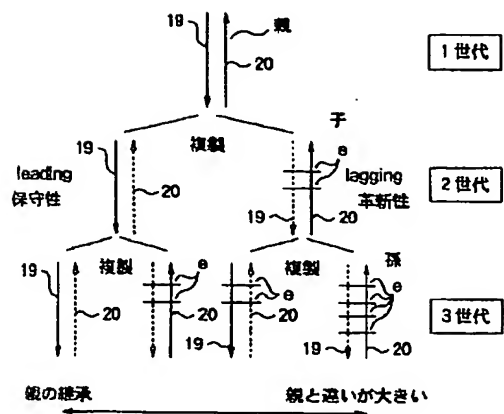
【図1】



【図2】



【図3】



【図4】

